

The *SM 22* promoter directs tissue-specific expression in arterial but not in venous or visceral smooth muscle cells in transgenic mice

H. Moessler¹, M. Mericskay², Z. Li², S. Nagl, D. Paulin² and J. V. Small^{1,*}

¹Institute of Molecular Biology, Austrian Academy of Sciences, Billrothstrasse 11, A-5020 Salzburg, Austria

²Biologie Moleculaire de la Differentiation, Universite Paris 7 et Institut Pasteur SCME 25, rue du Dr Roux, 75724 Paris, France

*Author for correspondence

SUMMARY

The transcriptional signals underlying smooth muscle differentiation are currently unknown. We report here the complete sequence and characterization of the single mouse gene for the smooth muscle-specific protein SM 22 and the transcriptional activity of its promoter in cultured smooth muscle cells in vitro and in transgenic mice. In the transgenic animals, promoter constructs ranging in length from 445 to 2126 bp directed reporter expression initially in the heart and the somites of embryos and subsequently in the arteries of the vascular system, but in none of the visceral muscles, nor in the veins. Expression in the heart was spatially restricted to the presumptive right ventricle and outflow tract and disappeared in the adult. Likewise, expression in the somites was only transitory and was not

observed after about 14.5 days post coitum in the embryo. In the adult mouse, *SM 22* promoter activity persisted in the smooth muscle cells of the arteries and was still notably absent from other smooth muscles, despite the ubiquitous presence of the endogenous SM 22 protein. These findings on the transcriptional activity of a smooth muscle promoter in vivo reveal the existence of different differentiation programmes for smooth muscle cells in the veins and the arteries and raise the expectation of a further subdivision of programmes among the visceral muscles.

Key words: *SM 22*, promoter, transgenic mice, smooth muscle cells, developmental expression

INTRODUCTION

The discovery of specific transcription factors that programme progenitor cells along the myogenic pathway has provided a new level of understanding of muscle differentiation (Olson and Klein, 1994; Buckingham, 1994), but which is mainly restricted to skeletal muscle. Despite much effort, matching progress has not been made with cardiac muscle and even less with smooth muscle. In particular, the absence of an in vitro model of smooth muscle differentiation has precluded the use of strategies so successfully employed with skeletal muscle (Davis et al., 1987) to uncover determinants of the smooth muscle lineage. An alternative, but less direct, approach involves initial characterization of promoters of genes encoding products specific for the smooth muscle phenotype, and entailing expression mapping of putative regulatory elements in vivo, in transgenic mice. The present report concerns such an analysis with the gene encoding the smooth muscle-specific protein SM 22.

SM 22 is an abundant protein in smooth muscles, first isolated from chicken gizzard and bovine aorta by Lees-Miller et al. (1987a,b) and sequenced by the same group (Pearlstone et al., 1987). Subsequently, SM 22 has reappeared and been cloned in various guises (p27; WS3-10; transgelin) according to its noted expression in different in vitro systems, in particular in fibroblasts, where it can be upregulated either by serum stimulation (Almendral et al., 1989) or ageing (Thweatt et al.,

1992) or down-regulated by viral transformation (Shapland et al., 1993; Prinjha et al., 1994). In contrast to this promiscuous expression in cells in vitro, the in vivo expression of SM 22 faithfully follows that of other smooth muscle markers such as calponin (Gimona et al., 1990; Frid et al., 1992) both during embryogenesis and in the adult (Duband et al., 1993; Lees-Miller et al., 1987b). Like calponin, SM 22 has also been singled out as a differentiation marker of smooth muscle via differential cDNA screening between cultivated primary smooth muscle cells and the same cells after multiple passages (Shanahan et al., 1993).

As compared to other products of the differentiated smooth muscle cell, SM 22 offers particular advantages as a phenotypic marker. It is uniformly expressed in smooth muscle unlike the smooth muscle actin isoforms (Vandekerckhove and Weber, 1981) and does not exhibit alternatively spliced variants as arise from the genes encoding caldesmon (Hayashi et al., 1992), metavinculin (Koteliansky et al., 1992) and the myosin heavy chain (Babji and Periasamy, 1989; Nagai et al., 1989), which complicate analysis of expression patterns. Finally, its small size makes SM 22 more amenable to cloning strategies.

For the foregoing reasons, the attention of several groups, including our own (Solway et al., 1995; Kemp et al., 1995; Osbourn et al., 1995; Li et al., 1996a, Mössler, 1995) has recently turned to the analysis of the expression of the *SM 22* gene to unveil smooth muscle-specific promoter elements. By

extending the analysis to expression in transgenic mice, we have been able to delineate a regulatory region including -445 bp upstream of the transcriptional start site that directs specific expression in arterial smooth muscles *in vivo*. More surprisingly, the findings reveal that a specific regulatory programme is required for different smooth muscles, since the transgene expression in our lines is tightly restricted to the arterial system, being excluded even from veins, although SM 22 is expressed in both vascular and visceral muscles. While this work was undergoing revision, a report by Olson and coworkers appeared (Li et al., 1996b) describing very similar results.

MATERIALS AND METHODS

Cloning and characterization of the murine SM 22 gene

Approximately 10^6 recombinant phage from a murine C57/bl Lambda EMBL3 genomic library (kindly provided by Dr D. Plachov, Munster, Germany) were screened with the 890 bp *BaII-EcoRV* murine SM 22 cDNA probe (Almendral et al., 1989; generously provided by Dr R. Bravo) labelled by random priming with α -[32 P]dCTP. Hybridization was performed overnight at 65°C in Church buffer. The filters were subsequently washed in 0.2× SSC (citrate-buffered saline), 0.1% SDS (sodium dodecyl sulfate) at 65°C and positive plaques were purified to homogeneity through three successive rounds of rescreening under identical conditions. Three overlapping clones were isolated and one of them was found to contain the entire SM 22 coding region and was further used for restriction mapping and subcloning. The absence of rearrangements during the cloning procedure was confirmed by Southern blot analysis of mouse genomic DNA and of the genomic clones. The sequence of the SM 22 gene was determined on both strands by the dideoxy chain-termination method using the Sequenase V2.0 Kit (United States Biochemical, Cleveland, Ohio).

Primer extension, 5'RACE and RNase protection analysis

Total RNA from cell lines and adult mouse tissue was isolated by the method of Chirgwin et al. (1979). mRNA was isolated with the Oligo-Tex mRNA purification kit (Quiagen Inc., Chatsworth, CA) according to the manufacturer's instructions. For primer extension analysis a 30-mer oligonucleotide designed to include the reverse complement of base pairs +36 to +65 of the SM 22 cDNA (5'-AAGGCTTG-GTCGTTTGTGGACTGGAAGGAG-3') was 5'-labelled with T4-poly-nucleotide kinase. Mouse uterine mRNA (5 µg) was hybridized at 55°C to 10^5 cts/minute of probe and the primer extension reaction was carried out according to standard laboratory procedures (Maniatis et al., 1982). The reaction products were analysed by electrophoresis on 6% denaturing polyacrylamide gels. A sequence reaction using the same primer was loaded in the adjacent lanes to enable a direct determination of the transcriptional initiation site.

5'-RACE was performed using two 18 bp antisense oligonucleotides corresponding to base pairs +112 to +129 (5'-TTTGTG-GACTGCATTCTC-3') and +162 to +179 (5'-AATCCACTCCAC-TAGTCG-3') of the SM 22 cDNA, respectively, essentially as described by Frohman (1990). For RNase protection analysis, a 417 bp PCR fragment spanning the region from -352 to +65 of the SM 22 gene was produced and blunt-end cloned into the *HincII* site in a pBluescript SK+ vector. The sequence of the PCR fragment was verified. A α - 32 P-labelled antisense RNA probe was generated from this construct linearized with *SphI* at position -203 of the SM 22 gene with T7 RNA polymerase. The probe (3×10^5 cts/minute) was hybridized to 5 µg of mouse uterine mRNA at 42°C overnight. After hybridization, samples were incubated with RNase A and T (1 hour at 37°C), precipitated and analysed as described for primer extension analysis.

Reporter constructs

Constructs used for transient expression experiments and transgenic animals were produced in a multistep process. First, a pGEM 7 (Promega Inc., Madison, WI) plasmid was digested with *NaeI/Bsp120I*, the *Bsp120I* overhang of the vector was filled-in with Klenow enzyme and the plasmid was religated in the presence of a 10 bp *SalI* linker, thereby deleting the *lacZ* fragment of pGEM7, introducing a *SalI* site and restoring the *Bsp120I* site. Subsequently, a *PmeI* site was introduced into the polylinker by ligating a *PmeI* linker with single-stranded *NsiI* overhangs into the *NsiI* site of the construct. Finally, the *lacZ* (β -galactosidase) gene coupled to a nuclear localization signal was isolated from pDes2.nlz (Li et al., 1993) with *HindIII* and *SacI* and inserted between the *HindIII* and the *SacI* site of the modified pGEM 7 vector. This procedure yielded the promoterless vector pnlz. Construct p352nlz, which comprises the proximal 352 bp of the SM 22 promoter together with the first exon, was made by inserting the 417 bp PCR fragment mentioned above into the *XbaI* and *XhoI* sites of modified pGEM 7 vector. To construct p2126nlz, a 1.9 kbp *Bsp120I/SphI* extending from position -2126 to position -206 of the SM 22 promoter was ligated into p352nlz, which was digested with *XbaI/SphI* to remove the fragment -352 to -206. Plasmid p445nlz was generated from p2126nlz by removing the fragment from -2126 to -445 by digestion of *ApaI* and *PstI*. Plasmid p2126INTnlz was constructed by inserting a 5.8 kbp *Bsp120I/HindIII* fragment spanning the region from bp -2126 to 3651 of the SM 22 locus into pnlz, giving pINT-nlz. Subsequently, a 496 bp PCR fragment spanning from bp 3648 to 4143 of the SM 22 locus was generated using a perfectly matching upstream primer and antisense downstream primer, which was designed to incorporate an artificial *HindIII* site at position 4135-4140. This PCR product was cut with *HindIII* and ligated into the *HindIII* site of pINT-nlz to yield p2126INTnlz. and correct orientation of the insert was verified. The construct p2126INTnlz therefore contained 2126 bp of proximal upstream region, the first exon, the entire first intron together with the first 15 bp of the second exon but with the last 4 bp mutated from CATG to GCTT to delete the translational start codon and to introduce a *HindIII* site as outlined above.

Cell culture and transfection

Cultured rabbit aortic cells (J. Daniel-Laregière and J. Bonnet, personal communication) were grown under two conditions: in the proliferative state in the presence of 10% serum or without serum. In the absence of serum, all cells stained with anti-smooth muscle actin and desmin antibodies, indicating upregulation of smooth muscle phenotypic markers in this condition. For transfection experiments, rabbit aorta cells were seeded into 6 cm Petri dishes and grown to approximately 75% confluence before adding the transfection mixture. Transient transfections were performed using the Lipofectamine reagent (Life Technologies Inc.). In brief, 24 µg of Lipofectamine was added to 6 µg of supercoiled reporter gene plasmid. After adding 3 ml of DMEM the transfection cocktail was added to the cells. This mix was replaced after 6 hours with medium with or without 10% serum. Expression of the reporter gene product was analysed after 48 hours by measuring β -galactosidase activity using the Dual-Light system (TROPIX Inc., Bedford, Massachusetts) according to manufacturer's instructions. Relative activities were calculated as the mean of two independent experiments. The mean activity of the pRSV *lacZ* was taken as 100%.

Production of transgenic mice

Inserts of p445nlz, p2126nlz and p2126INTnlz were isolated by digestion with *SalI/PmeI* and subsequent gel electrophoresis, further purified using the GeneClean kit (Bio101, La Jolla, CA) and finally resuspended in 10 mM Tris-HCl pH 7.5, 0.1 mM EDTA. Transgenic animals were produced essentially as described previously (Li et al., 1993), by injecting DNA into pronuclei of fertilized oocytes of inbred FVB mice. Eggs that survived injection were transplanted into the

oviduct of pseudopregnant B6/CBAF1 fosters. Embryos were allowed to go to term and transgene expression was analyzed in the F₁ embryos obtained by timed mating. Mice carrying the transgene were identified by PCR and Southern blotting using tail DNA. For staging embryos, the day that the plug was observed was considered to be 0.5 days post coitum (d.p.c.). Construct 2126nlz gave rise to three positive animals that transmitted and expressed the transgene. For construct 2126INTnlz, two positive animals out of 28 were obtained and two expressed the construct. For experimental matings, transgenic mice were backcrossed with C57/bl mice and all histochemical staining experiments were performed with mice hemizygous for the transgene.

Alternatively, (for construct p445nlz) founder embryos were killed at 13.5 d.p.c. and analysed for transgene expression. Three positive embryos expressing this transgene were obtained.

Histochemical staining

Whole embryos or adult tissues were fixed for 15-30 minutes at 4°C in 1% formaldehyde in buffer A (100 mM sodium phosphate buffer pH 7.3, 2 mM MgCl₂, 0.01% sodium deoxycholate and 0.02% NP-40). After rinsing, specimens were stained overnight at 30°C in buffer A containing 1 mg/ml X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside; Sigma, St Louis, Missouri), 5 mM potassium ferrocyanide, 5 mM potassium ferricyanide, and 20 mM Tris-HCl pH 7.3. Stained samples were rinsed again, dehydrated through increasing concentrations of ethanol, cleared in xylene and embedded in paraffin. Histological sections were cut at a thickness of 7-10 µm and counterstained with Ehrlich's hematoxylin and eosin. For whole-mount observation, embryos were cleared in a mixture of benzyl alcohol and benzyl benzoate (Dr E. Rovin, personal communication).

For immunohistochemical staining, mouse tissues were fixed in 3% paraformaldehyde in phosphate-buffered saline (PBS) at 4°C for 3 hours, embedded in Tissue-tek medium (Leica, Vienna) and sectioned at 5-10 µm on a Reichert cryomicrotome (Leica, Vienna). Selected sections were mounted and stained for SM 22 using a monoclonal antibody against SM 22 (Duband et al., 1993) at a dilution of 1/30 and a Cy3-labelled secondary antibody (Biological Detection Systems, USA). Sections were viewed on a Zeiss inverted 135 TV fluorescence microscope and images collected via a back-cooled CCD camera (Princeton Instruments, Princeton, USA).

RESULTS

Characterization of the SM 22 locus

Hybridisation screening of a mouse genomic library with the SM 22 cDNA probe yielded three overlapping clones, one of which contained the complete SM 22 gene. A restriction map of this clone, which is partially depicted in Fig. 1, was established. The complete sequence of the SM 22 gene was determined on both strands and is shown in Fig. 2. Exons were located by comparison with the SM 22 cDNA sequence (Almendral et al., 1989; Solway et al., 1995). The gene (EMBL database, accession No. Z68618) spans 5923 bp from the transcriptional start site to the site of polyadenylation and the coding part is split in 5 exons (Fig. 1). All exon-intron boundaries match the consensus sequences (AG/GT). The first exon contains only leading sequence (65 bp) and the translational start codon is therefore located in the second exon separated from the translational start site by more than 4 kb of intronic sequence. Two putative polyadenylation signals are located at position 5905 and 5913 and several TGT-islands occur downstream of the polyadenylation site,

which presumably take part in the termination of transcription (Birnstiel et al., 1985). Sequence analysis further revealed the presence of three repeat sequences (Fig. 1) highly homologous to B1 elements.

The transcriptional start site was mapped by primer extension analysis (Fig. 3). An antisense oligonucleotide complementary to the region +36 to +65 of exon 1 resulted in four extension products, the longest of which was 65 bp, indicating the start of the transcript at a G, 28 bp downstream of the putative TATA box. The slightly shorter extension products were most likely caused by premature termination of the primer extension reaction due to modifications (capping) at the 5'-end of the mRNA. The assignment of the start site was confirmed by RNase protection analysis yielding a protected fragment of 65 bp and by cloning and sequencing of a primer extension product derived from an independent reaction with an antisense primer located in the second exon of the gene (not shown).

The sequence TTATAA at position -28 bp (Fig. 2) is closely related to the consensus TATAAA sequence and presumably functions as a TATA-box. Computational sequence analysis of the 5'-upstream region revealed the presence of many putative binding sites for ubiquitous transcription factors as well as elements that are known to contribute to the transcriptional regulation of muscle genes (Fig. 2). We found a total of 11 E-boxes (CANNTG; Murre et al., 1989), four Mef-2/rSRF motifs (YTAWAAATAR; Gosset et al., 1989), four putative SRF-binding elements (CC(A/T)6GG; Gustafson et al., 1988), five AP-2-binding sites (CCCMNSSF; Mitchell et al., 1987), and five SP 1 core motifs (GGGCGG; Dynan and Tjian, 1983). Finally, we could locate an element homologous to the TGT3-3 motif (McNamara et al., 1995), which was recently shown to bind a smooth muscle nuclear factor and to contribute to the transcriptional regulation of the smooth muscle-specific α-actin gene.

Expression in vitro

In a first attempt to identify the *cis*-acting DNA sequences involved in the regulation of the SM 22 gene, a series of constructs were tested in transient transfection experiments. Four constructs (p352nlz, p445nlz, p2126nlz, p2126INTnlz) were prepared using the bacterial *lacZ* gene with the nuclear localization signal of SV40 virus as reporter (Fig. 4A). The construct p352nlz and p445nlz started respectively from -352 or -445 bp upstream of the initiation site and extended to 65 bp in the first exon. The construct p2126nlz contained 2126 bp of proximal upstream region plus the first exon (65 bp) and the p2126INTnlz was identical to p2126nlz except for the addition of the first intron and the first 12 bp of the second exon (Fig.

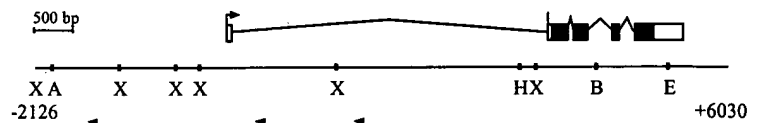


Fig. 1. General features of the SM 22 gene. Restriction map and exon/intron structure of the SM 22 locus embracing the region -2126 to +6030. Transcriptional start site is indicated by right-angled arrow. Exons are boxed, with open boxes corresponding to non-coding regions. Arrowheads indicate positions of B1-repeats (see Fig. 2). Abbreviations of restriction sites: A, *Apa*I; B, *Bam*HI; E, *Eco*RV; H, *Hind*III; X, *Xba*I.

[illegible]

[illegible]

4A). The expression of these four constructs was tested under two conditions in cultured rabbit aortic cells: in the proliferative state (in the presence of 10% serum) or after serum withdrawal, under which condition markers of the differentiated

phenotype have been shown to be upregulated. After transfection of these different constructs, β -galactosidase was assayed using the x-gal reaction. As shown in Fig. 4B, comparison of transient expression of *lacZ* demonstrated that the construct

containing 2126 bp conferred high level expression in the absence of serum, whereas a 3-fold lower expression was found when cells were cultured with 10% serum. In the absence of serum, construct p445nlz showed a similar activity to p2126nlz, suggesting that this shorter region was sufficient to direct transcription. No expression was detected with the p206nlz recombinant. In a previous attempt to identify the *cis*-acting DNA sequences involved in the regulation of the *SM 22* gene, Solway et al. (1995) also found that the 441 bp region of the *SM 22* 5'-flanking sequence was necessary and sufficient to program high level transcription of a luciferase reporter gene in both primary rat aortic SMCs and A7r5 cells in vitro.

Expression in transgenic mice

To determine whether the promoter constructs derived from the *SM 22* gene direct appropriate expression smooth muscle cells in vivo, we generated transgenic mice harboring the different *SM 22-nls/lacZ* transgenes. For in vivo analysis, we first used the construct p2126nlz to test if this 5' flanking region contained sufficient information to target smooth muscle-specific expression of the transgene. We obtained three positive animals that transmitted and expressed the transgene. Two independent lines expressing this transgene were used for our initial studies. Embryos from 8.5 to 15.5 d.p.c. were stained for β -gal activity. As shown in Fig. 5C-E, primary and progressive expression occurred in the vascular system of the developing embryo and was accompanied by transient expression (see further below) in the heart and somites.

Specific expression of the *SM 22-lacZ* transgene in arterial but not in venous or visceral smooth muscle cells during embryogenesis

In the vascular system, expression was detected first in the dorsal aorta at around 9.5 d.p.c. (Fig. 5B), progressing at 10.5 d.p.c. to the aortic arches and at 11.5 d.p.c. (Fig. 5C) to the iliac arteries, both umbilical arteries, the carotid arteries and the major vessels of the head. Transgene expression in the vasculature further increased through 12.5 d.p.c., at which time the intercostal vessels could readily be seen (Fig. 5D) and by 15.5 d.p.c. (Fig. 5E) β -gal was present in all the major vessels including those in the limbs and the tail. At this time, the staining of the pulmonary trunk also became evident. Histological sections revealed unexpected differences in the expression of the transgene in arteries and veins. The section through a 12.5 d.p.c. embryo in Fig. 6A shows high levels of expression in the muscle layers of both carotid arteries, the fourth aortic arch arteries, the proximal pulmonary trunk and the proximal part of the ascending aorta, whereas both the left and right anterior cardinal veins are unstained. The lack of transgene expression in venous muscle was also evident in sections through the ductus venosus, as well as the portal and the umbilical vein (not shown). The same situation was observed in the adult animal, where expression could be detected neither in the

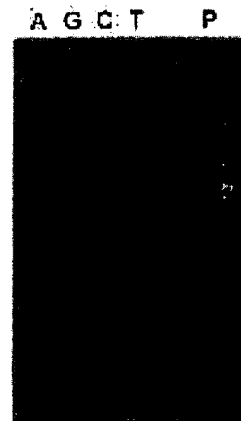


Fig. 3. Mapping of the transcription initiation site of the mouse *SM 22* gene by primer extension. Lane P, primer + 5 μ g mouse uterine mRNA; lane A, C, G, T dideoxy sequence ladder; sequence reaction was done by using the same primer. Transcription initiation site is indicated by arrowhead.

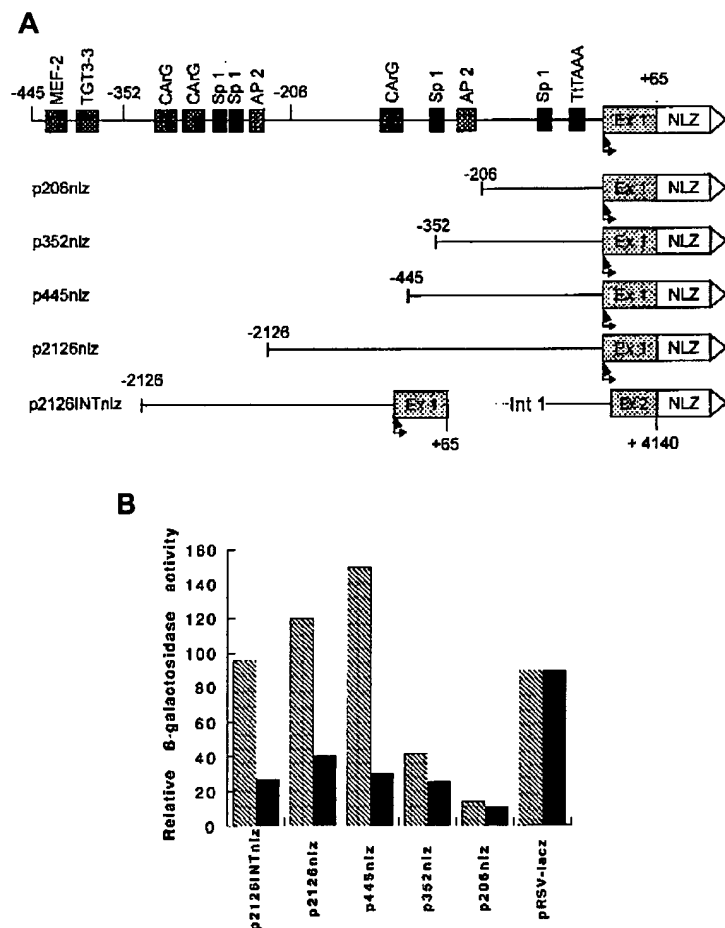


Fig. 4. (A) Schematic representation of constructs of the *SM 22* gene using bacterial β -galactosidase (nlz) as reporter. Summary of the putative transcription factor-binding sites in the proximal promoter region are indicated. Ex 1 and Ex 2 indicate incorporated exons and Int 1, intron 1. Numbers indicate positions relative to the transcriptional start point (right-angled arrow). (B) Transient transfection analyses of the *SM 22*nlz reporter plasmids in the cultured aortic cells under proliferating conditions (solid bars) and after serum withdrawal (differentiated conditions; hatched bars). Activities for each plasmid were determined as described in Methods.

vena cava (Fig. 7A) nor the pulmonary veins (not shown). These observations pointed to previously unrecognized differences in the smooth muscle layers of arteries and veins. Higher magnification views of arteries showed that expression was restricted to the muscle layer and was absent from the endothelium (Fig. 6C).

It became evident from the whole-mount preparations (Fig. 5D,E) that transgene expression was absent from visceral smooth muscle tissue. This was unexpected, as the endogenous *SM 22* gene is known to be expressed in vascular and visceral smooth muscle tissue (see Introduction). In histological sections through the abdominal region of a 14.5 d.p.c. embryo, the muscle layers of stomach, gut and the developing bladder are easily recognized (Fig. 6B). Expression of the transgene in this region was clearly detected in both umbilical arteries, but in the visceral muscle layers of the stomach, the duodenum, the midgut, the hindgut, and the bladder essentially no β -gal expression was seen. Furthermore, expression was also absent from the esophagus, the trachea (Fig. 6A), and the bronchi of the lung (not shown).

In view of the fact that Solway et al. (1995) had shown that a region embracing 441 bp was sufficient to programme high level transcription in smooth muscle cells *in vitro*, we analysed the expression of the shorter construct p445nlz containing 445 bp of 5'-flanking sequence to +65 bp in exon 1. The three founder embryos (13.5 d.p.c.) obtained with this transgene all showed the same expression pattern as constructs p2126nlz and p2126INTnlz (not shown), with the exception of one, which was particularly strongly stained, that showed an additional, but weak ectopic expression in the ribs. These results demonstrated that the regulatory elements sufficient for specific expression of the *SM 22* gene in the arterial smooth muscle cell *in vivo* are located between -445 and +65 bp.

Transient expression of *SM 22-lacZ* in the right ventricle of the heart and somites

As shown in Fig. 5, expression was initially detected at 8.5 d.p.c. in the heart region of the embryo (Fig. 5A) and was also seen in the vessels of the surrounding yolk sac (not shown). At 9.5 d.p.c., expression of *lacZ* in the heart increased and was confined to the presumptive right ventricle and the outflow tract (Fig. 5B). With further development from 11.5 to 15.5 d.p.c., a regional segregation of label became evident and was confined to the right ventricle (Fig. 5C,D). Sections showed that expression occurred throughout the myocardium of the right ventricle with some minor scattered label in the left ventricle and the right atrium (Fig. 6A). This expression in the heart was only transitory, being absent in the adult (not shown). Early transgene expression was also observed in the rostral somites starting with 9.5 d.p.c. (Fig. 5B), extending to more caudal somites at 10.5 d.p.c. (not shown) and reaching a peak at 11.5 d.p.c. (Fig. 5C). Expression then diminished, starting with the rostral somites at 12.5 d.p.c. (Fig. 5D) and finally was no longer detectable after 14.5 d.p.c. (Fig. 5E). Expression in each somite lasted about 2 days, and could be seen in sections to be restricted to the myotomal region of the somites (Fig. 6D).

Restricted expression persisted in the adult

The expression in the arterial system persisted in the adult, being most clearly visible in the aorta, the pulmonary trunk and

the right pulmonary artery (Fig. 7A), as well as in vessels of the intestine (Fig. 7B), the bladder and the uterus (not shown).

Fig. 7B shows that transgene expression was absent in the muscular wall of the colon, whereas the adjacent mesentery vessel was strongly positive for β -gal. Endogenous *SM 22* protein could however be demonstrated in both tissues by immunofluorescence microscopy, using an antibody against *SM 22* (Fig. 7C). The same antibody also labelled smooth muscle cells in both veins and arteries in the adult animal (Fig. 7D). The lack of transgene expression in adult visceral smooth muscle was further confirmed by staining specimens of the esophagus, the trachea, the bladder, the vas deferens and the uterus for β -gal activity. In no case could any expression in the muscle layers of these tissues be detected (not shown).

Addition of intron 1 (p2126INTnlz) does not modify expression pattern

Since a large intron within the 5'-leading sequence seems to be a common feature of smooth muscle-specific genes (Reddy et al., 1990; Babij et al., 1991; Miwa et al., 1991; Mössler, 1995; Solway et al., 1995; Kemp et al., 1995), we speculated that the 4 kb long intron 1 sequence could be involved in the regulation of the *SM 22* gene in other smooth muscle tissues. Using the longer construct (p2126INTnlz), we obtained two positive animals out of 28, one of which transmitted and expressed the construct. The expression pattern of p2126INTnlz containing intron 1 was identical to p2126nlz at all stages observed, except that the staining intensity was slightly less due either to a lower copy number of the transgene or to a low splicing efficiency of the primary transcript. This result suggests that the 4 kb long intron 1 is not essential for *SM 22*-specific expression.

DISCUSSION

Features of the *SM 22* promoter

The expression of *SM 22* is directed from a single copy gene now cloned from both mouse (Solway et al., 1995; Mössler, 1995) and rat (Osbourne et al., 1995; Kemp et al., 1995). Comparison of the reported mouse sequence with our own reveal some minor base differences, presumably due to strain differences; larger deviations from that of the rat are however evident. Osbourne et al. (1995) have identified an extra intron and two transcripts for the rat gene, but neither northern blots (Mössler, 1995) nor the present primer extension assays gave any evidence for this in mouse (see also Solway et al., 1995). Common features of the mouse and rat proximal promoters include two CARG boxes (Gustafson et al., 1988) in similar positions and several SP 1 motifs. The serum activation of *SM 22* expression in non-muscle cells (Mössler, 1995; Almendral et al., 1989) is similar to that described for α -smooth muscle actin (Kim et al., 1993) and presumably involves the CARG motifs. For α -smooth actin, these motifs endow some degree of smooth muscle specificity *in vitro* and bind nuclear proteins from cultured smooth muscle cells (Shimizu et al., 1995; Cogan et al., 1995; Blank et al., 1995). CARG motifs are also present in the smooth muscle myosin heavy chain gene (Katoh et al., 1994) but their wide distribution in genes for cardiac and skeletal muscle (Sartorelli et al., 1990; Lee et al., 1991) indicate that they are not smooth muscle specific. Nevertheless, reporter

assays with cultured smooth muscle cells have shown that truncation of the mouse *SM 22* promoter to exclude the more distal CARG motif (at -273) reduced expression by 80% (Solway et al., 1995) although a less dramatic effect was observed for a similar construct of the rat *SM 22* promoter (Kemp et al., 1995).

MADS-Box transcription factors belonging to the MEF-2 family have been shown to be expressed in smooth muscle tissue (Yu et al., 1992). Recently, it has been found that the single MEF-2 isoform in *Drosophila* is required for the formation of all three muscle types (Bour et al., 1995; Lilly et al., 1995). Five A/T-rich elements matching the MEF-2 consensus sequence are present in the regulatory region of the *SM 22* gene. It remains to be elucidated whether these elements are binding sites for MEF-2 isoforms and take part in the regulatory process of the *SM 22* gene.

A total of 11 E-boxes could be found in the upstream region of the *SM 22* gene, although in the proximal promoter region (-445 bp) no E-box is present. The activity of the *SM 22* regulatory region seems to be unaffected by these elements both in vitro in cultured smooth muscle cells (Solway et al., 1995) and in vivo (our results), suggesting that basic helix-loop-helix (bHLH) transcription factors are not required for the expression of smooth muscle-specific genes. According to the current paradigm in developmental biology (Olson, 1990), additional factors and their binding motifs restricted to the smooth muscle lineage are required for the specific regulatory programme in these cells. A possible candidate is the TGT3-3 motif, which has recently been shown to confer positive regulation to the smooth muscle α -actin gene in cultured smooth muscle cells and to bind nuclear factors extracted from them (McNamara et al., 1995). Significantly this motif also features in the *SM 22* genes of both mouse and rat (this study; Solway et al., 1995; Kemp et al., 1995).

Restricted expression of the *SM 22* transgene in vivo

The transient expression that we observed for all constructs in the heart region and in the somites of the embryo correspond to that described for the *SM 22* endogene using in situ hybridization (Li et al., 1996a). The transient expression in the progenitor cells of skeletal and cardiac muscle is not surprising in view of the promiscuous expression of smooth muscle actin isoforms in these muscles during development (McQuinn and Schwartz, 1995). This overlap in expression of gene products also occurs between skeletal and cardiac muscle (Kelly et al., 1995). Most interesting in the latter respect is that the transient expression of the skeletal muscle myosin light chain gene (encoding MLC3F) is spatially restricted to the atria and the left ventricular compartments of the embryonic mouse heart in a comple-

mentary fashion to the *SM 22-lacZ* gene expression, which occurs in the outflow tract and the presumptive right ventricle. Our findings thus provide further support for the existence of differentially programmed cell populations in different regions of the heart (Kelly et al., 1995).

More strikingly, the present results demonstrate that there is a subdivision of expression programmes for smooth muscle cells according to which organs they invest. Had this subdivision been between visceral and vascular muscles it would have been less surprising, but we see that even the smooth muscle progenitors of the arteries and veins are directed by different developmental cues. The findings in this report point to differences in smooth muscle cell populations that were not formerly recognized.

Tracing the lineage of smooth muscle differentiation has been hampered by the lack of early markers (Owens, 1995) and, in consequence, little has been established about the

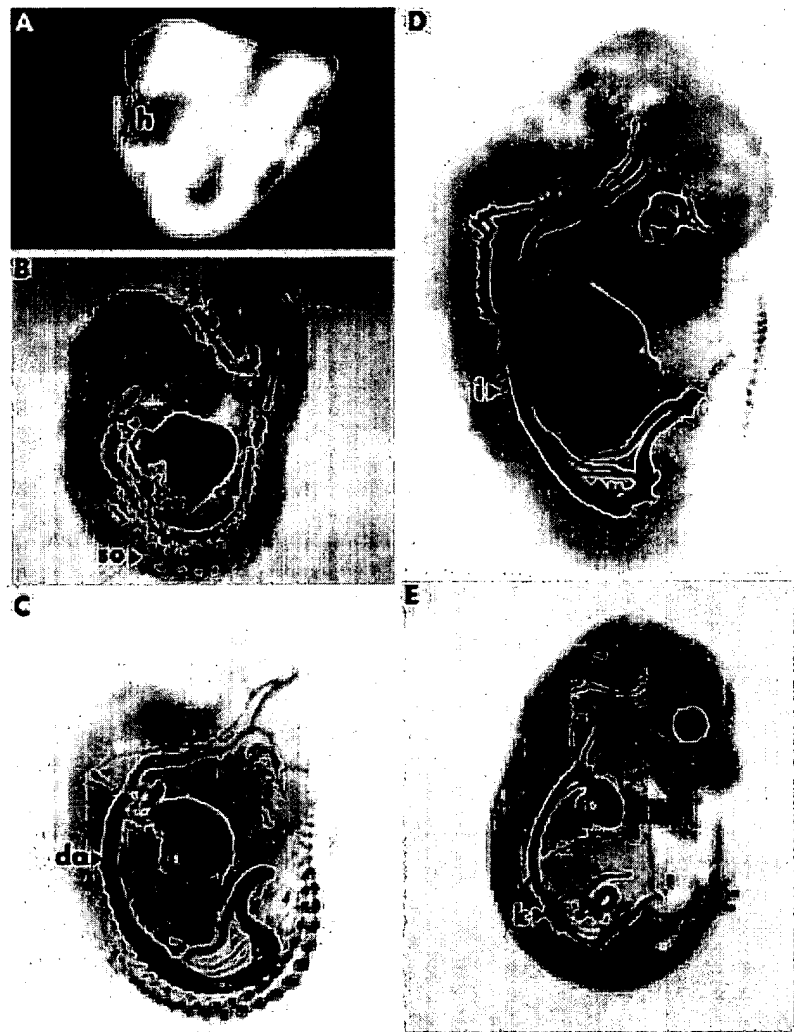


Fig. 5. *SM 22* expression in transgenic mouse embryos. Whole-mount pictures of β -gal-stained embryos of the following ages: 8.5 d.p.c. (A), 9.5 d.p.c. (B), 11.5 d.p.c. (C), 12.5 d.p.c. (D) and 15.5 d.p.c. (E). For details see text. Abbreviations: h, heart region; so, somites; da, dorsal aorta; i, intercostal vessels; k, kidney.

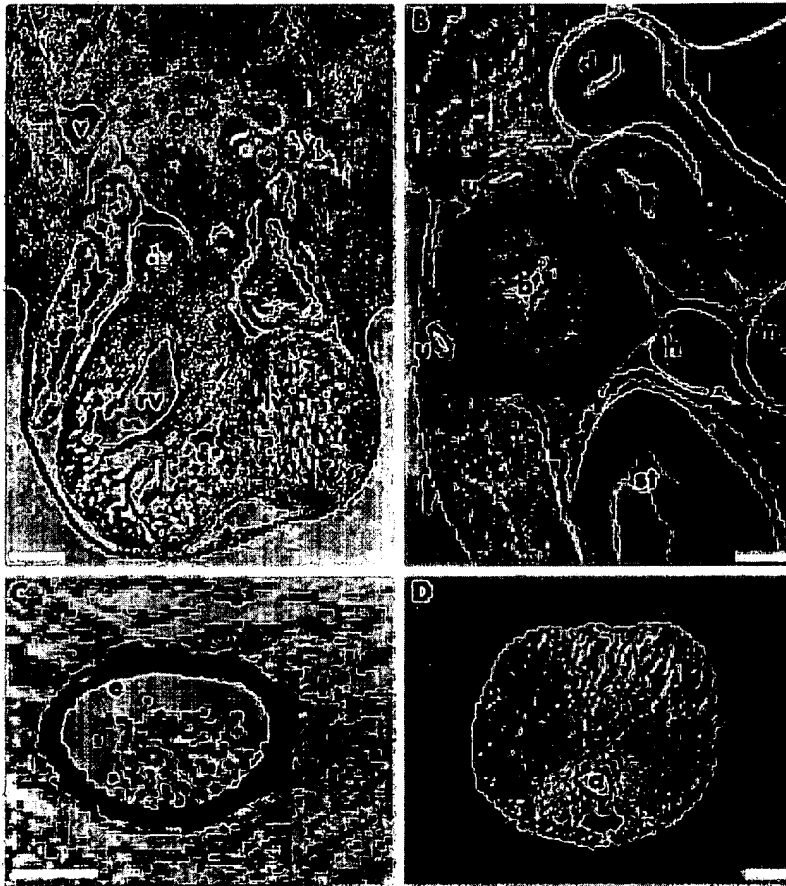


Fig. 6. Histological sections of β -gal-stained embryos. (A) Section through a 12.5 d.p.c. embryo at the level of the heart showing intense staining exclusively in the myocardium of the right ventricle (rv) and in the walls of the arteries. Abbreviations: a, proximal part of the aorta; a', fourth aortic arch arteries; c, carotid arteries; e, esophagus; lv, left ventricle; p, pulmonary trunk; rv, right ventricle; t, trachea; v, left and right cardinal vein. Bar, 200 μ m. (B) Section through the abdominal region of a 14.5 d.p.c. embryo showing staining of both umbilical arteries (u) and absence of label in the visceral organs: b, bladder; d, duodenum; h, hindgut; l, liver; m, midgut; st, stomach. Bar, 200 μ m. (C) Higher magnification image of umbilical artery of a 12.5 d.p.c. embryo showing exclusive labelling of muscle layer and negative reaction in the endothelium (e). Bar, 50 μ m. (D) Section through a tail somite (12.5 d.p.c.) showing label in the myotome and the artery of the tail (a). Bar, 100 μ m.

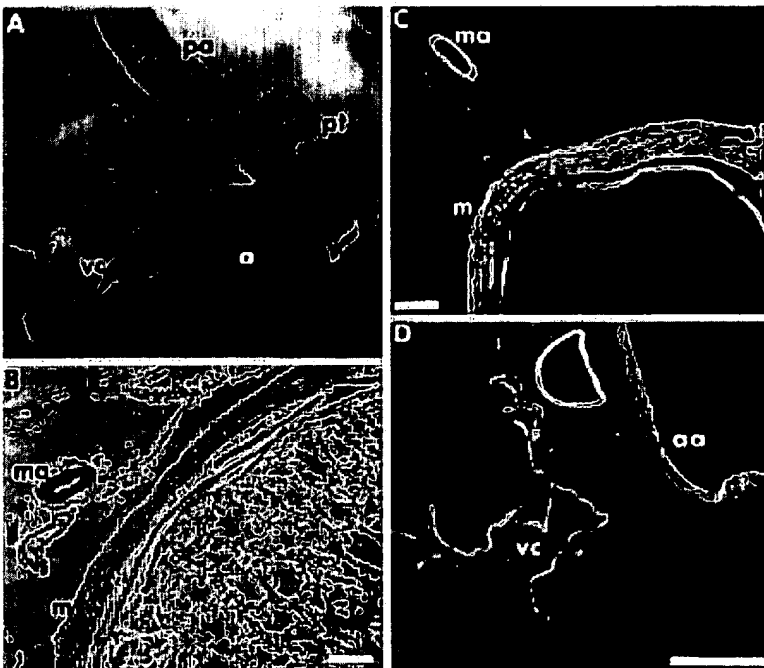


Fig. 7. Transgene expression in adult mice (6-8 weeks old). (A) Top view of a heart whole mount showing intense staining of aorta (a), the pulmonary trunk (pt), and the pulmonary artery (pa) and absence of staining in the vena cava (vc). (B,C) Sections through the smooth muscle layer (m) of the adult colon and an adjacent mesentery artery (ma). Staining for β -gal (B) showed that transgene expression is restricted to the artery, whereas immunofluorescence with the SM 22 antibody (C) revealed expression of endogenous SM 22 in smooth muscle cells of both artery and colon. Bars, 100 μ m. (D) Section through abdominal aorta (aa) and vena cava (vc) of an adult animal labelled with the SM 22 antibody and showing SM 22 protein expression in both vessel types. Bar, 500 μ m.

origins of smooth muscle cells, apart from the observation that different populations arise from the mesoderm and the neural crest (Gittenberger de Groot et al., 1995). Since smooth muscle cells readily de-differentiate in vitro (see above) and show the same plasticity in vivo, especially in response to lesions in the arterial wall (Ross, 1993), we can expect that extrinsic positional cues (Owens, 1995) play an important part in signalling and maintaining the differentiated smooth muscle phenotype. Smooth muscle cell layers are typically formed around the epithelial borders of the organs requiring their contractile properties, so that inductive cues, if they exist, most likely derive from the epithelium. The type of smooth muscle cell programme set in action could then be defined by the character of the epithelium, giving the possibility for subtle modulations in the properties of the associated smooth muscle cells. In this context, we would expect the existence of additional elements in the *SM 22* gene that respond to different extrinsic transcriptional cues. Taking this line of argument, the present promoter construct may include elements responsive only to cues from the arterial system. In view of the recent findings showing that the expression of the muscle protein desmin in skeletal and cardiac muscle (Li et al., 1993; and unpublished results), and by inference in smooth muscle, involves spatially different regions of the single desmin gene, we might expect that different regions in the *SM 22* gene specify expression in different smooth muscles.

In conclusion, the present findings raise new and interesting questions about the factors that determine a subdivision of the smooth muscle system as well as about its functional and physiological consequences. Further, the identification of a promoter region specific for expression in the arterial wall raises interesting possibilities of applying gene therapy following arterial injury (Ohno et al., 1994). After completion and submission of this work, a paper by Li et al. (1996b) appeared, describing very similar findings. We confirm and extend the observation of these authors.

We thank Mrs M. Schmittner for excellent technical assistance and Mr A. Weber for photographic assistance. We also thank Professor K. Kratochwil for his advice and for critically reading the manuscript. Transgenic mice used in this study were obtained from the Service facility of Dr I. Cerutti (SEAT/IFC-1) CNRS, Villejuif. This work was supported by grants from the Austrian Science Research Council and in part by grants from the Association française contre les myopathies (AFM) and a contract from Ministère de l'Education Nationale et de la recherche (MRES) no. 1087.

REFERENCES

- Almendral, J. M., Santaren, J. F., Perera, J., Zerial, M. and Bravo, R. (1989). Expression, cloning and cDNA sequence of a fibroblast serum-regulated gene encoding a putative actin-associated protein (p27). *Exp. Cell Res.* **181**, 518-530.
- Babij, P. and Periasamy, M. (1989). Myosin heavy chain isoform diversity in smooth muscle is produced by differential RNA processing. *J. Mol. Biol.* **210**, 673-679.
- Babij, P., Kelly, C. and Periasamy, M. (1991). Characterisation of a mammalian smooth muscle myosin heavy-chain gene: Complete nucleotide and protein coding sequence and analysis of the 5' end of the gene. *Proc. Natl. Acad. Sci. USA* **88**, 10676-10680.
- Birnstiel, M. L., Busslinger, M. and Strub, K. (1985). Transcription termination and 3' processing: the end is in site. *Cell* **41**, 349-359.
- Blank, R. S., McQuinn, T. C., Yin, K. C., Thompson, M. M., Takeyasu, K., Schwartz, R. J. and Owens, G. K. (1995). Elements of the smooth muscle α -actin promoter required in cis for transcriptional activation in smooth muscle. *J. Biol. Chem.* **267**, 984-989.
- Bour, B. A., O'Brien, M. A., Lockwood, W. L., Goldstein, E. S., Bodmer, R., Taghert, P. H., Abmayr, S. M. and Nguyen, H. T. (1995). Drosophila MEF2, a transcription factor that is essential for myogenesis. *Genes Dev.* **9**, 730-741.
- Buckingham, M. (1994). Molecular biology of muscle development. *Cell* **78**, 15-21.
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. and Rutter, W. J. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* **18**, 5294-5299.
- Cogan, J. G., Sun, S., Stoffet, E. S., Schmidt, L. J., Getz, M. J. and Strauch, A. R. (1995). Plasticity of vascular smooth muscle α -actin gene transcription. *J. Biol. Chem.* **270**, 11310-11321.
- Davis, R. L., Weintraub, H. and Lassar, A. B. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* **51**, 987-1000.
- Duband, J.-L., Gimona, M., Scatena, M., Sartore, S. and Small, J. V. (1993). Calponin and SM 22 as differentiation markers of smooth muscle: spatiotemporal distribution during avian embryonic development. *Differentiation* **55**, 1-11.
- Dynan, W. S. and Tjian, R. (1983). The promoter-specific transcription factor SP 1 binds to upstream sequences in the SV 40 early promoter. *Cell* **35**, 79-87.
- Frid, M. G., Shekhonin, B. V., Koteliensky, V. E. and Glukhova, M. A. (1992). Phenotypic changes of human smooth muscle cells during development: Late expression of heavy caldesmon and calponin. *Dev. Biol.* **153**, 185-193.
- Frohman, M. A. (1990). RACE: Rapid amplification of cDNA ends. In *PCR Protocols: A Guide to Methods and Applications* (ed. M. A. Innis, D. H. Gelfand, J. J. Sninsky and T. J. White), pp 28-38. San Diego: Academic Press Inc.
- Gimona, M., Herzog, M., Vandekerckhove, J. and Small, J. V. (1990). Smooth muscle specific expression of calponin. *FEBS Lett.* **274**, 159-162.
- Gittenberger-De Groot, A. C., Slomp, J., Deruiter, M. C. and Poelman, R. E. (1995). Smooth muscle cell differentiation during early development and during intimal thickening formation in the ductus arteriosus. In *The Vascular Smooth Muscle Cell* (ed. S. M. Schwartz and R. P. Mecham), pp 17-36. San Diego: Academic Press Inc.
- Gosset, L. A., Kelvin, D. J., Sternberg, E. A. and Olson, E. N. (1989). A new myocyte specific enhancer-binding factor that recognizes a conserved element associated with multiple muscle-specific genes. *Mol. Cell. Biol.* **9**, 5022-5033.
- Gustafson, T. A., Miwa, T., Boxer, L. M. and Kedes, L. (1988). Interaction of nuclear proteins with muscle-specific regulatory sequences of the human cardiac α -actin promoter. *Mol. Cell. Biol.* **8**, 4110-4119.
- Hayashi, K., Yano, H., Hashida, T., Takeuchi, R., Takeda, O., Asada, K., Takahashi, E., Kato, I. and Sobue, K. (1992). Genomic structure of the human caldesmon gene. *Proc. Natl. Acad. Sci. USA* **89**, 12122-12126.
- Katoh, Y., Loukianov, E., Kopras, E., Zilberman, A. and Periasamy, M. (1994). Identification of functional promoter elements in the rabbit smooth muscle myosin heavy chain gene. *J. Biol. Chem.* **269**, 30538-30545.
- Kelly, R., Alonso, S., Tajbakhsh, S., Cossu, G. and Buckingham, M. (1995). Myosin light chain 3F regulatory sequences confer regionalized cardiac and skeletal muscle expression in transgenic mice. *J. Cell Biol.* **129**, 383-396.
- Kemp, P. R., Osbourn, J. K., Grainger, D. J. and Metcalfe, J. C. (1995). Cloning and analysis of the promoter region of the rat SM 22 α gene. *Biochem. J.* **310**, 1037-1043.
- Kim, J.-H., Bushel, P. R. and Kumar, C. C. (1993). Smooth muscle α -actin promoter activity is induced by serum stimulation of fibroblast cells. *Biochem. Biophys. Res. Com.* **190**, 1115-1121.
- Koteliensky, V. E., Ogryzko, E. P., Zhidkova, N. I., Weller, P. A., Critchley, D. R., Vancompernelle, K., Vandekerckhove, J., Strasser, P., Way, M., Gimona, M. and Small, J. V. (1992). An additional exon in the human vinculin gene specifically encodes meta-vinculin-specific difference peptide. *Eur. J. Biochem.* **204**, 767-772.
- Lee, T.-C., Chow, K.-L., Fang, P. and Schwartz, R. J. (1991). Activation of skeletal α -actin gene transcription: The cooperative formation of serum response factor-binding complexes over positive cis-acting promoter serum response elements displaces a negative acting nuclear factor enriched in replicating myoblasts and nonmyogenic cells. *Mol. Cell. Biol.* **11**, 5090-5100.
- Lees-Miller, J. P., Heeley, D. H., Smillie, L. B. and Kay, C. M. (1987a).

- Isolation and characterization of an abundant and novel 22-kDa protein (SM 22) from chicken gizzard smooth muscle. *J. Biol. Chem.* **262**, 2988-2993.
- Lees-Miller, J. P., Heeley, D. H. and Smillie, L. B. (1987b). An abundant and novel protein of 22 kDa (SM 22) is widely distributed in smooth muscles. *Biochem. J.* **244**, 705-709.
- Li, L., Miano, J. M., Cserjesi P. and Olson, E. N. (1996a). SM22 α , a marker of adult smooth muscle, is expressed in multiple myogenic lineages during embryogenesis. *Circ. Res.* **78**, 188-195.
- Li, L., Miano, J. M., Mercer, B. and Olson, E. N. (1996b). Expression of the SM22 α promoter in transgenic mice provides evidence for distinct transcriptional regulatory programs in vascular and visceral smooth muscle cells. *J. Cell Biol.* **132**, 849-859.
- Li, Z., Marchand, P., Humbert, J., Babinet, C. and Paulin, D. (1993). Desmin sequence elements regulating skeletal muscle-specific expression in transgenic mice. *Development* **117**, 947-959.
- Lilly, B., Zhao, B., Ranganayakulu, G., Paterson, B. M., Schulz, R. A. and Olson, E. N. (1995). Requirement for MADS domain transcription factor D-MEF2 for muscle formation in *Drosophila*. *Science* **267**, 688-693.
- Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982). *Molecular Cloning*. New York: Cold Spring Harbour.
- McNamara, C. A., Thompson, M. M., Vernon, S. M., Shimizu, R. T., Blank, R. S. and Owens, G. K. (1995). Nuclear proteins bind a cis-acting element in the smooth muscle α -actin promoter. *Am. J. Physiol.* **268**, 1259-1266.
- McQuinn, T. C. and Schwartz, R. J. (1995). Vascular smooth muscle specific gene expression. In *The Vascular Smooth Muscle Cell* (ed. S. M. Schwartz and R. P. Mecham), pp. 213-261, San Diego: Academic Press Inc.
- Mitchell, P. J., Wang, C. and Tjian, R. (1987). Positive and negative regulation of transcription in vitro: Enhancer-binding protein AP-2 is inhibited by SV40 T-antigen. *Cell* **50**, 847-861.
- Miwa, T., Manabe, Y., Kurokawa, K., Kamada, S., Kanda, N., Bruns, G., Ueyama, H. and Kakunaga, T. (1991). Structure, chromosome location and expression of the human smooth muscle (enteric type) γ -actin gene: Evolution of six human actin genes. *Mol. Cell. Biol.* **11**, 3296-3306.
- Mössler, H. (1995). Characterization of the gene structure and the promoter region of the smooth muscle specific protein SM 22 of *mus musculus*. PhD thesis at the University of Salzburg, Austria.
- Murre, C., McCaw, P. S., Vässin, H., Caudy, M., Jan, L. Y., Jan, Y. N., Cabrera, C. V., Buskin, J. M., Hauschka, S. D., Lassar, A. B., Weintraub, H. and Baltimore, D. (1989). Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* **58**, 537-544.
- Nagai, R., Makoto, K., Babij, P. and Periasamy, M. (1989). Identification of two types of smooth muscle myosin heavy chain isoforms by cDNA cloning and immunoblot analysis. *J. Biol. Chem.* **264**, 9734-9737.
- Ohno, T., Gordon, D., San, H., Pompili, V. J., Imperiale, M. J., Nabel, G. J. and Nabel, E. G. (1994). Gene therapy for vascular smooth muscle cell proliferation after arterial injury. *Science* **265**, 781-784.
- Olson, E. N. (1990). Myo D family: a paradigm for development?. *Genes Dev.* **4**, 1454-1461.
- Olson, E. N. and Klein, W. H. (1994). bHLH factors in muscle development: dead lines and commitments, what to leave in and what to leave out. *Genes Dev.* **8**, 1-8.
- Osbourne, J. K., Weissberg, P. L. and Shanahan, C. M. (1995). A regulatory element downstream of the rat SM 22 α gene transcription start point enhances reporter gene expression in vascular smooth muscle cells. *Gene* **154**, 249-253.
- Owens, G. K. (1995). Molecular identity of smooth muscle cells: Overview. In *The Vascular Smooth Muscle Cell* (ed. S. M. Schwartz and R. P. Mecham), pp. 163-167. San Diego: Academic Press Inc.
- Pearlstone, J. R., Weber, M., Lees-Miller, J. P., Carpenter, M. R. and Smillie, L. B. (1987). Amino acid sequence of chicken gizzard smooth muscle SM 22 α . *J. Biol. Chem.* **262**, 5985-5991.
- Prinjha, R. K., Shapland, C. E., Hsuan, J. J., Totty, N. F., Mason, I. J. and Lawson, D. (1994). Cloning and sequencing of cDNAs encoding the actin cross-linking protein transgelin defines a new family of actin-associated proteins. *Cell Motil. Cytoskeleton* **28**, 243-255.
- Reddy, S., Ozgur, K., Lu, M., Chang, W., Mohan, S., Kumar, C. and Ruley, H. E. (1990). Structure of the human smooth muscle α -actin gene. *J. Biol. Chem.* **265**, 1683-1687.
- Ross, R. (1993). The pathogenesis of atherosclerosis: a perspective for the 1990s. *Nature* **362**, 801-809.
- Sartorelli, V., Webster, K. A. and Kedes, L. (1990). Muscle-specific expression of the cardiac α -actin gene requires MyoD1, CArG-box binding factor, and Sp1. *Genes Dev.* **4**, 1811-1822.
- Shanahan, C. M., Weissberg, P. L. and Metcalf, J. C. (1993). Isolation of gene markers of differentiated and proliferating vascular smooth muscle cells. *Circ. Res.* **73**, 193-204.
- Shapland, C., Hsuan, J. J., Totty, N. F. and Lawson, D. (1993). Purification and properties of Transgelin: A transformation and shape change sensitive actin-gelling protein. *J. Cell Biol.* **121**, 1065-1073.
- Shimizu, R. T., Blank, R. S., Jervis, R., Lawrenz-Smith, S. C. and Owens, G. K. (1995). The smooth muscle α -actin gene promoter is differentially regulated in smooth muscle versus non-smooth muscle cells. *J. Biol. Chem.* **270**, 7631-7643.
- Solway, J., Seltzer, J., Samaha, F. F., Kim, S., Alger, L. E., Niu, Q., Morrissey, E. E., Ip, H. S. and Parmacek, M. S. (1995). Structure and expression of a smooth muscle cell-specific gene, SM 22 α . *J. Biol. Chem.* **270**, 13460-13469.
- Thweatt, R., Lumpkin, C. K. and Goldstein, S. (1992). A novel gene encoding a smooth muscle protein is overexpressed in senescent human fibroblasts. *Biochem. Biophys. Res. Com.* **187**, 1-7.
- Vandekerckhove, J. and Weber, K. (1981). Actin typing on total cellular extracts: Highly sensitive protein-chemical procedure able to distinguish different actins. *FEBS Lett.* **113**, 595-603.
- Yu, Y.-T., Breitbart, R. E., Smoot, L. B., Lee, Y., Mahdavi, V. and Nadal-Ginard, B. (1992). Human myocyte-specific enhancer-factor 2 comprises a group of tissue-restricted MADS box transcription factors. *Genes Dev.* **6**, 1783-1798.

(Accepted 13 May 1996)